

УДК 004.032

**Володимир Красиленко, Діана Нікітович**  
**Вінницький національний технічний університет**  
**Вінниця, Україна**

**АНАЛІЗ ДИНАМІКИ НАУКОВОЇ ПУБЛІКАЦІЙНОЇ  
ДІЯЛЬНОСТІ КРАЇН СВІТУ ЗА ЇХ ІНТЕГРАЛЬНИМИ  
ПОКАЗНИКАМИ ТА ІНДИКАТОРАМИ ЦИТУВАННЯ В  
СВІТОВІЙ СИСТЕМІ SJR**

***Анотація:** Наведені результати порівняльного аналізу стану та динаміки наукової публікаційної активності та результативності країн світу на основі інтегральних щорічних показників та індикаторів цитування з системи SJR.*

***Ключові слова:** інтелектуальний аналіз, наукова публікаційна діяльність, наукометричні показники, індекси цитування, рейтингове оцінювання, h-індекс.*

***Аннотация:** Представлен анализ состояния и динамики научной публикационной активности и результативности стран мира на основе интегральных ежегодных SJR показателей и индикаторов цитирования.*

***Ключевые слова:** научная публикационная деятельность, анализ, наукометрические показатели, индекс цитирования, рейтинговое оценивание.*

***Abstract:** The analysis of the state and dynamics of scientific publication activity and effectiveness of the countries of the world is presented on the basis of integral annual SJR indicators and citation indicators.*

***Keywords:** citation indicators, cluster analysis, scientific publication activity.*

**Вступ, огляд публікацій.** Для порівняльного аналізу результативності наукової публікаційної діяльності (НПД) у різних країнах світу та наукових напрямках чи їх галузях, визначення для них рейтингових місць, часток їх вкладу у цілісну картину світового наукового середовища, надання світовій спільноті інформації створені

та розвиваються світові системи, наприклад, Scimago Journal & Country Rank (SJR) [1] та їхні національні складові [2].

Оскільки одним з основних наукових продуктів є публікація (стаття в журналі, матеріали чи тези доповіді на конференції, монографія, патент, тощо), а для об'єктивного наукового оцінювання результативності, продуктивності НПД використовують низку наукометричних показників (П): кількість публікацій (документів), індекси цитування, імпакт-фактори, як числові показники рівня наукового журналу та статей у ньому, то для забезпечення загальних та часткових потреб створені, створюються та функціонують універсальні, регіональні, дисциплінарні науко-метричні бази даних, серед універсальних такі як, Scopus, Web of Knowledge, Index Copernicus, Google Scholar, ResearchGate, та інші, а також представлені в Science Citation Index (SCI) і Journal Citation Reports (JCR) цитатні бази, що випускаються інститутом ISI у Філадельфії, США, та містять бібліографічні описи всіх статей з оброблюваних наукових журналів і відображають в основному публікації з фундаментальних розділів науки в провідних міжнародних і національних журналах. На основі показників цих баз визначають рейтинги науковців, країн, ВНЗ, установ, міністерств, наукових фахових видань, тощо [2].

Проте у всіх рейтингових таблицях, так і в SJR місця країн, журналів чи науковців ранг визначається за одним індикатором або послідовно за іншими при однаковості попередніх.

*У SJR країни ранжуються за показником загальної кількості наукових публікацій за визначений період років, а для експертного оцінювання результативності діяльності вчених і дослідницьких колективів використовуються індекс Гірша чи йому подібні.*

Але, хоч набір таких індикаторів суттєво і розширився, особливо в останні роки, представлені рейтинги базуються у вищевказаних системах на консолідованих абсолютних а не відносних показниках та не враховують їх взаємозалежність, впливовість на результуючу оцінку чи місце. А це призводить до необ'єктивного оцінювання, викривлення результатів і висновків. Вирішення цих проблем запропоновано нами у роботах [4-6] шляхом використання інтелектуальних інформаційних технологій (ІІТ), нейропакетів Excel

Neural Package, Neuro Pro для багатоаспектного групування та інтегрального визначення сукупного рейтингу країн, України на основі їх показників, індексів у SJR стосовно їх НПД.

Переваги нашого підходу, що дозволив ефективно здійснювати кластерний аналіз об'єктів із значною кількістю якісних та кількісних ознак і визначати впливовість факторів, були ефективно продемонстровані на обробці реальних даних, взятих з SJR та інших наукометричних баз, створили підґрунтя для прийняття оптимальних стратегій стосовно розвитку аналогічних систем в Україні так і науки в цілому та дали більш точний інструмент для висновків та пропозицій щодо покращення НПД та її продуктивності. Кластерному аналізу 200 країн світу за показниками їх НПД та індикаторів цитування, що представлені в SJR, присвячена робота [6], проте динаміка рейтингів країн та цих показників, особливо з урахуванням появи даних за останні роки, не досліджувалась.

**Тому метою даної роботи** є вивчення та аналіз динаміки НПД країн та її результативності для основних наукових галузей за період 2006-2016 роки на основі даних з SJR (за 2017 р. дані відсутні!).

**Об'єктом дослідження** є показники з SJR (станом на 01.03.2017) для вибраних країн, що взяті нами для експериментів, аналізу і демонстрації, такі як: загальна кількість (К) документів (Д), К цитованих Д, К цитувань та К самоцитувань, Н-індекс, та додатково розрахованих нами: К Д на 1 тисячу населення, середню кількість цитувань документа та відносні частки цитованих документів та самоцитувань.

**Предметом** дослідження є уточнення змін у кластерах групувань, аналіз НПД і показ переваг застосування ІТ для визначення рейтингів країн та їх динаміки, виявлення тенденцій у динаміці показників НПД України та країн з лідируючих кластерів, пояснення їх причин та шляхів покращення наукових здобутків.

**Виклад основного матеріалу та результатів аналізу.** Вивчення та порівняння статистичних даних з SJR показало, що крім нововведених 2016 р., показники за попередні роки (станом на кінець 2015 р.), які і були використані у [6], зазнали коригування в сторону їх збільшення, що обумовлено коректними причинами часових затримок

процесів відстеження, збору та обробки даних. Ці зміни для всіх країн, важливих з точки зору мети та аспектів дослідження, виявились незначними для 1, 2, 5 (Н-індекс, принципово є кумулятивним!), 6 (К населення), 9 (частка цитованих Д), 10 (частка самоцитувань) показників (П). Середні значення відхилень за взяті останні 11 років всіх цих П майже для всіх країн не перевищують 2-4%, а коригування спричинило суттєві зміни лише 3-го (К цитувань), 4-го (К самоцитувань), 7-го (середня К цитувань Д), наприклад, для України приріст 3-го відповідно за 2006-2015 склав: 15, 17, 19, 22, 27, 35, 40, 67, 142, 575 %; приріст 4-го теж відповідно склав: 11,4, 12,7, 15,8, 18,5, 23,1, 29,1, 36,7, 62,8, 130,1, 516,1 %; а приріст 7-го: 13,9, 16,3, 17,9, 22,0, 26,3, 25,2, 39,3, 65,7, 136,9, 476,9 %.

Аналіз показав, що навіть ці зміни не призвели до відчутних змін 9-го (середнє відхилення на рівні сотих 1%, а максимальне 1,6% у 2015 р., абсолютне його значення коливається у межах від 93% до 99% з середнім 97%), 10-го П (частка самоцитувань за вищевказані роки зменшилась в середньому на 3%, хоч для більшості років, крім останнього (на 8,8%), вона зменшувалась на 2-4%, а її значення коливаються в межах від 24% до 33% з середнім 28%, але намітилась тенденція до її збільшення, хоч на нашу думку об'єктивні причини для цього відсутні!).

Таким чином, можна зробити наступні **висновки**:

1) На 7 з 10 вибраних П щорічні коригування відчутних впливів не дають,

2) Аналіз динаміки НПД доцільно робити окремо для груп показників, мало змінних (1-2, 6, 8-10) та змінних (3, 4, 5, 7) від коригувань, при цьому для першої групи бажано не враховувати 1 чи 2 останні роки, а для другої більш динамічної (кумулятивної) або враховувати значний вплив коригувань або виключати навіть більше 5 останніх років;

3) Оскільки на 70% вибраних з 10 П індикаторів для розбиття країн на кластери коригування не впливають, то, як і підтверджують наші експерименти, розподіл країн на 11 груп на уточнених інтегральних даних (станом на 01.03.2017) практично не відрізняється

від їх розподілу, отриманого за допомогою нейро-пакетів у [6], особливо для лідируючих угруповань;

4) Тому для спрощення порівняльний аналіз динаміки НПД можна робити лише для окремих країн з типових вибраних представників кожного кластеру; 5) Виявлення річних тенденцій, трендів можливе, як традиційними статистичними методами та інструментами, так і за допомогою ІТ та нейрон-пакетів.

Результати першого експерименту, в якому проводився поділ країн на 11 кластерів на уточнених даних за допомогою карт Кохонена нейропакету Excel Neural Package [3], у вигляді фрагменту результуючої таблиці (вікна пакету) з номерами і показниками визначених кластерів та їх місць показані на рис. 1 та у табл. 1, де номери кластерів впорядковані у відповідності до отриманої К балів та наведені їх центроїдні характеристики. Оскільки всі П є кількісними, то дані після округлень використовувались безпосередньо без спеціального кодування.

32	Monaco	1 436	1 322	24 908	2 882	69	31	19	47	92	12 A	3
33					Кластер A10							
34	Russian Federation	701 029	689 095	4 289 618	1 273 073	390	142 424	6	5	98	30 A	10
35	Brazil	598 234	573 988	5 036 027	1 699 530	379	204 260	9	3	96	34 A	10
36	Taiwan	491 560	477 442	4 790 230	1 075 153	331	23 415	10	21	97	22 A	10
37	Poland	431 016	418 917	3 491 958	901 545	371	38 562	8	11	97	26 A	10
38	Turkey	390 874	368 197	2 938 841	737 423	266	79 414	8	5	94	25 A	10
39	Iran	287 010	278 388	1 504 541	573 856	180	81 824	5	4	97	38 A	10
40	Czech Republic	213 209	206 738	1 867 611	440 161	294	10 645	9	20	97	24 A	10
41	Mexico	210 387	201 540	1 962 572	408 914	289	121 737	10	2	96	21 A	10
42	Portugal	189 052	179 134	2 096 242	407 892	297	10 825	12	17	95	19 A	10
43	Malaysia	153 378	148 844	670 387	183 198	165	30 514	5	5	97	27 A	10
44	Argentina	145 416	138 788	1 681 700	354 132	273	43 432	12	3	95	21 A	10
45	Hungary	136 034	130 299	1 660 840	264 809	301	9 898	13	14	96	16 A	10
46	Ukraine	133 650	131 490	635 570	176 428	174	44 429	5	3	98	28 A	10
47	Romania	125 576	122 884	619 956	153 395	167	21 666	5	6	98	25 A	10
48	Egypt	120 493	117 104	818 728	162 544	165	88 487	7	1	97	20 A	10
49	Chile	90 216	86 521	1 014 687	193 534	233	17 508	12	5	96	19 A	10
50	Pakistan	81 612	78 219	425 467	118 262	148	199 086	5	0	96	28 A	10
51	Slovakia	72 847	70 964	561 511	113 733	180	5 445	8	13	97	20 A	10
52	Croatia	72 110	69 675	459 356	97 694	177	4 465	7	16	97	21 A	10
53	Slovenia	64 483	62 170	611 672	115 616	189	1 983	10	33	96	19 A	10
54	Nigeria	53 298	51 223	272 400	61 408	115	181 562	5	0	96	23 A	10
55	Tunisia	51 590	49 230	276 247	60 183	109	11 037	6	5	95	22 A	10
56	Serbia	45 000	43 151	188 381	47 922	100	7 177	4	6	96	25 A	10
57	Lithuania	32 137	31 399	227 339	51 689	133	3 526	7	9	98	23 A	10
58	Estonia	25 458	24 479	313 735	55 005	162	1 265	13	20	96	18 A	10
59					Кластер A9							
60	Thailand	109 832	104 982	976 328	162 255	213	67 976	9	2	96	17 A	9
61	Saudi Arabia	61 460	57 643	547 187	90 163	164	17 543	8	1	96	18 A	9

Рис.1. Результати кластеризації (фрагмент вікна програми з виділеними в кластер A10 країнами (Україна !).

Центр кластеру	Кластери										
	A1	A3	A2	A4	A10(5)	A7(6)	A8(7)	A11(8)	A6(9)	A5(10)	A9(11)
X <sub>1</sub> - Documents	8 626 193	282 994	1 439 678	1 850	196 627	4 015	2 255	2 307 950	485	4 414	19 066
X <sub>2</sub> - Citable documents	7 876 234	265 435	1 352 490	1 734	189 995	3 729	2 155	2 257 142	405	4 089	18 362
X <sub>3</sub> - Citations	177 434 935	4 760 553	22 005 038	46 847	1 536 625	32 490	30 301	13 049 752	4 273	59 019	133 919
X <sub>4</sub> - Self-Citations	83 777 658	731 220	5 231 040	4 541	389 084	3 499	2 985	6 435 573	272	7 591	21 467
X <sub>5</sub> - H index	1 648	412	739	83,2	218	48,9	50,5	438	26,8	67	82,4
X <sub>6</sub> - Population (in thousands)	321 369	7 035	68 089	1 855	55 384	4 996	8 503	1 310 178	7 967	21 978	29 757
X <sub>7</sub> - V Citations per Document	22,1	17,9	16,4	27,3	7,99	8,76	12,9	6,38	8,67	14,8	6,62
X <sub>8</sub> - Number of documents per 1000 population	26,3	40,1	24,6	2,7	9,19	2,87	1,72	1,72	2,06	1,07	1,68
X <sub>9</sub> - Fraction Citations	91,5	93,2	93,5	93,2	96,5	92,6	95,8	96,6	81,8	92,2	96,5
X <sub>10</sub> - Fraction Self-Citations	47,3	14,7	23,4	11,3	23,5	8,24	7,72	44,5	6,07	8,61	14,5
К-ть елементів в кластері	1	18	7	4	25	42	34	2	6	22	39
оцінка (бали)	11	10	9	8	7(2)	6(5)	5(3)	4	3(6)	2(7)	1
Бали_Міся(7_8)	10+10	11+9	8+9	11+6	8+3	7+5	6+4	1+3	4+5	7+1	2+1
Бали Міс (7_8_5)	20+11	20+8	17+10	17+6	11+7	12+3	10+3	4+9	9+1	8+4	3+5

Табл. 1 Характеристика виявлених кластерів

З показаних на рис. 2-3 ( послідовно: 1-4, 7, 9-10) П та результатів їх обробки встановлено, що починаючи з 2010 р. намітились майже стійкі позитивні тенденції до росту 1 і 2 П, 3-4, 7 (з урахуванням майбутніх позитивних коригувань) та 10-го П, проте з 2014 р. цей ріст призупинився і триває без змін.

	Country (Ukraine)	Documents	Citable documents	Citations	Self-Citations	V_Citations per Document	Fraction Citations (%)	Fraction Self-Citations (%)
1								
2								
3	2006	6728	6653	52324	13355	8	99	26
4	2007	6930	6818	56626	13850	8	98	24
5	2008	7231	7130	53808	13510	8	99	25
6	2009	7069	6968	42962	12970	6	99	30
7	2010	7610	7446	42031	12100	6	98	29
8	2011	8286	8026	53286	13222	7	97	25
9	2012	9318	8705	47538	12418	5	93	26
10	2013	9956	9715	36413	11236	4	98	31
11	2014	10246	9990	30290	8851	3	98	29
12	2015	10220	9942	19813	5748	2	97	29
13	2016	10087	9605	6172	2014	1	95	33
14	min	6728	6653	6172	2014	1	93	24
15	mean	8516	8273	40115	10843	5	97	28
16	max	10246	9990	56626	13850	8	99	33
17	Summ	93681	90998	441263	119274	5	1070	307
18		8516	8273	40115	10843	5	97	28
19								

Рис.2. Результати ( вікно з уточненими даними для України, кластер A10!).

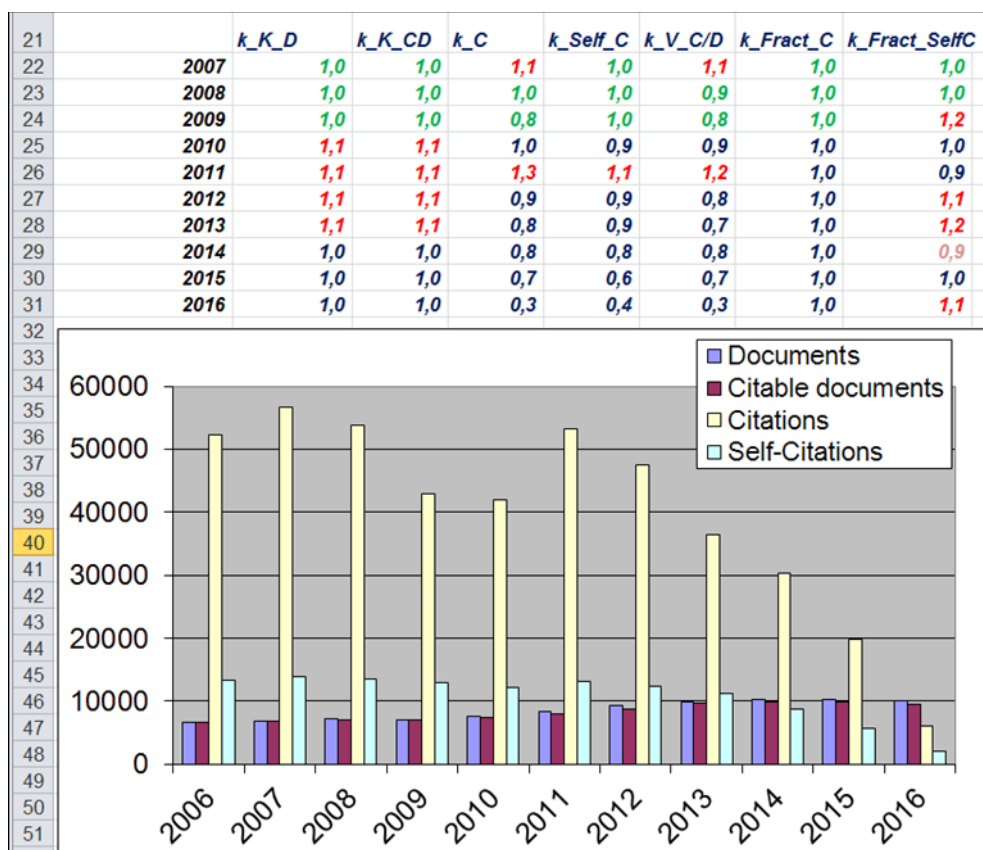


Рис.3. Вікно з діаграмою 4 П (Україна) та ланцюговими коефіцієнтами їх росту.

Аналогічно, можна зробити аналіз динаміки інших країн, що і буде предметом доповіді, а тут з урахуванням обмежень ми акцентуємо увагу лише на тому, що у США останні 8 років відсутні ріст та спад П НПД, наші сусіди Польща та Росія мають стале, хоч і незначне зростання, а Китай відзначається дуже позитивною динамікою (майже щорічним на 15-20% ростом П!).

**Висновок.** Дослідження показали, що лише інтелектуальний аналіз та групування дозволяють отримати більш точні обґрунтовані оцінки стану та динаміки активності, результативності НПД країн та їх місць в рейтингу.

### Література

1. Scimago Journal & Country Rank (SJR) [Електронний ресурс]. - Режим доступу: <http://www.scimagojr.com/countryrank.php>
2. Бібліометрика Української науки [Електронний ресурс]. - Режим доступу: [http://nbuviap.gov.ua/bpnu/index.php?page\\_sites=formy](http://nbuviap.gov.ua/bpnu/index.php?page_sites=formy)



3. Круглов В.В. Искусственные нейронные сети. Теория и практика / В.В. Круглов, В.В. Борисов. – М.: Горячая линия - Телеком, 2002. – 328 с.

4. Красиленко В.Г. Інтелектуальний кластерний аналіз наукометричних показників наукової діяльності кращих науковців планети з нейротематика за даними Google Scholar / В.Г. Красиленко, Д.В. Нікітович // Актуальні питання документознавства та інформаційної діяльності: Матеріали І МПК. – Одеса: ОНПУ, 2015. - С. 275-280.

5. Красиленко В.Г. Визначення впливовості факторів на інтегральну рейтингову оцінку кращих науковців світу з нейротематики на основі статистичних даних наукометричної бази Google Scholar / В.Г. Красиленко, Д.В. Нікітович // Актуальні питання документознавства та інформаційної діяльності: Матеріали І МПК. – Одеса: ОНПУ, 2015. - С. 333-338.

6. Красиленко В. Г. Кластерний аналіз країн світу за їх інтегральними за останні 20 років бібліометричними показниками наукових публікацій та цитувань / В. Г. Красиленко, Д. В. Нікітович // Інформаційна освіта та професійно-комунікативні технології ХХІ століття, Одеса, 2016. – С. 77-85. – Режим доступу: [http://opu-konf.at.ua/2016/krasilenko\\_v.g-nikitovich\\_d.v-klasternij\\_analiz\\_kr.pdf](http://opu-konf.at.ua/2016/krasilenko_v.g-nikitovich_d.v-klasternij_analiz_kr.pdf)